

Working group 2:
Data sharing & re-using

Stage 4b report:
**Deliver: What would make
the biggest impact?**

Last updated:
05 June 2025

Consultation window:
06 June -09 June 2025

EXECUTIVE SUMMARY

This Stage 4 report by Working Group 2 (WG2) of the Evidence Synthesis Infrastructure Collaborative (ESIC) outlines five high-priority strategies to advance the sharing and reuse of evidence synthesis data globally. These strategies are the culmination of extensive capability assessments, gap analysis, and multi-stage consultations across various regions, sectors, and stakeholder groups. The core aim is to normalize the practice of "studying a question once and reusing the answers across regions, languages, and sectors", thereby ensuring equity, interoperability, and sustainability within evidence systems. While each strategy addresses a distinct aspect of infrastructure, they are intrinsically linked and designed to function synergistically to realize this shared goal. A crucial cross-cutting strategy, Sustainable Funding and Incentive Alignment is foundational to the long-term viability and inclusive participation of all five technical strategies.

The prioritized strategies are:

- **Building a Federated Repository of Living Evidence Data:** This strategy proposes an interconnected system of evidence repositories, allowing each repository to maintain control over its data while contributing to broader reusability through shared standards and interoperability practices. Key features include federated nodes for sovereign or restricted data, multilingual access, AI-supported tagging and curation tools, and open APIs for cross-platform data integration. This architecture supports interoperability without requiring centralization.
- **Developing Standardized Record Structures and Interoperable Formats:** This strategy focuses on developing shared record structures for evidence synthesis inputs and outputs and adopting machine-readable data formats. This foundational layer aims to reduce redundancy, improve cross-platform data exchange, and enable automation, benefiting both current and legacy datasets through prospective and retrospective conversion.
- **Establishing Metadata standards to facilitate data identification and discoverability:** This strategy introduces modular metadata schemas and classification standards, including multilingual taxonomies and tagging protocols, to enhance evidence classification, searchability, and linkage across platforms. It emphasizes increasing the visibility of underrepresented outputs, such as non-English literature, LMIC evaluations, and grey literature, to foster greater equity and trust.
- **Advancing Open Access Standards for Equitable Sharing:** This strategy aims to establish open access policies and standards that ensure equitable sharing and reuse of evidence synthesis data by addressing legal, financial, and structural barriers. It promotes permissive licensing, ethical reuse, and legal clarity, and supports infrastructure for citation, versioning, and governance integration to recognize evidence as a global public good.
- **Ensuring Quality Assurance and Monitoring Systems:** This strategy embeds a tiered validation system alongside quality assurance (QA) and monitoring and evaluation (M&E) mechanisms within the infrastructure. It includes co-developed criteria for assessing completeness, relevance, reliability, and ethical compliance of shared data, and participatory review processes to ensure data integrity, equity, and system accountability. This strengthens user trust and system performance over time.

A critical enabler underpinning all five strategies is a cross-cutting approach to sustainable funding and incentive alignment. This includes establishing multi-donor pooled funds, aligning investment structures, and creating recognition and reward systems for data producers, reviewers, translators, and platform stewards. It also emphasizes equitable access to funding and support for LMIC institutions and under-recognized contributors, along with behavioural incentive models to encourage cultural shifts towards shared infrastructure stewardship.

The overall cost of all ESIC strategies is estimated at USD 56.2 million, with the five strategies in this group accounting for USD 55.98 million. This report will inform open consultation and final costing refinement, forming the basis for integration with other working groups in Stage 5. Many components—particularly those focused on metadata standards, record structures, open access, and quality mechanisms—are realistically implementable within 24–36 months, with early pilots possible in the first year. These strategies offer a credible, coordinated, and scalable pathway to make evidence reuse a global norm.

INTRODUCTION

This Stage 4 report presents the final strategies prioritized by Working Group 2 (WG2) of the Evidence Synthesis Infrastructure Collaborative (ESIC), focused on strengthening global systems for data sharing and reuse. It builds on previous phases that mapped capabilities (Stage 1), identified system gaps (Stage 2), and developed solutions (Stage 3), culminating in a set of strategies with the highest potential for scalable, equitable, and system-wide impact.

Through consultations and group deliberation, WG2 consolidated diverse proposals into five core strategies aimed at embedding evidence reuse into the fabric of global knowledge systems. These strategies reflect a shared goal: to conduct the core work of evidence synthesis once and reuse it many times across regions, languages, and sectors. Each strategy was assessed using an impact-effort lens and aligned with ESIC's principles of equity, interoperability, and sustainability.

The five strategies function as an integrated ecosystem. Standardized record structures (Strategy 2) enable metadata interoperability (Strategy 3), improving discoverability within a federated repository system (Strategy 1). Open access standards (Strategy 4) provide legal and policy foundations, while quality assurance mechanisms (Strategy 5) maintain trust, consistency, and integrity across the ecosystem. All are supported by a cross-cutting strategy on sustainable funding and incentives to ensure long-term participation and stewardship.

Recognizing institutional diversity and political realities, these strategies emphasize interoperability and modular adoption rather than rigid standardization. Co-designed protocols and opt-in mechanisms are central to enabling trust and alignment across systems with varying capacities and mandates.

APPROACH TO PRIORITIZE STRATEGIES AND SUMMARY OF RESULTS

A survey was conducted among WG members using a 5-point Likert-type scale. The survey included seven sections, each aligned with strategy groupings from Report #3. Members were asked to score each strategy in terms of impact and effort, and to provide a brief rationale for their scores.

An impact-effort matrix was generated from the survey results, using an average score of 3.5 as the threshold between high and low impact or effort (see Annex 2). Strategies were ranked by average scores (from high to low) and prioritized in batches as follows:

Batch 1 (broad consensus): Strategies with an interquartile range lower bound of 4.0 or higher on impact (i.e., at least 75% of respondents rated the strategy as high or extreme impact).

Batch 2 (moderate to high impact, low effort): Strategies with average impact scores above 3.5 and average effort scores below 3.5.

Strategies meeting these criteria were then grouped based on thematic overlap and complementarity.

Notably, “incentives to comply with existing infrastructure” was prioritized based on Working Group discussions, which identified behavior change as a critical enabler for uptake. While much of this report focuses on technical infrastructure, the WG emphasized that sustained use and alignment with shared systems will also require attention to cultural, institutional, and motivational incentives.

The final grouped strategies are presented below. Annex 3 outlines the relationship between these strategy groups and those included in the Stage 3 report. Annex 4 lists strategies that were not prioritized. Note that many strategies in Annex 4 overlap with those prioritized in this report and have been integrated accordingly.

PRIORITIZED STRATEGIES

Five strategies were finally prioritized and selected each contributing uniquely to the overarching vision of global evidence reuse:

- Building a Federated Repository of Living Evidence Data
- Developing Standardized Record Structures and Interoperable Formats
- Establishing Metadata standards to facilitate data identification and discoverability
- Advancing Open Access Standards for Equitable Data Sharing and Reuse
- Ensuring Quality and Monitoring of Data Sharing and Reuse Systems

While each strategy is presented separately for clarity, they are intended to operate as an integrated system. Several components (e.g., metadata, QA, access standards) are shared or cross-referenced across strategies, and their success

depends on coordinated implementation. The federated repository serves as a unifying layer that brings these technical and governance elements together in practice.

A sixth, cross-cutting strategy, Sustaining the Infrastructure: Funding Models and Incentives—was identified as essential to enable and support the long-term success of the five technical strategies. These strategies are not isolated but are designed to reinforce one another, with progress in one area often enabling advancements in others.

Each strategy will be supported by advisory input—either through dedicated or shared groups, depending on the implementation context. These groups will provide strategic oversight, guide early design, and ensure alignment with stakeholder expectations. A coordinated advisory model can enhance integration across strategies, particularly when partners are involved in multiple areas. Their role supports legitimacy, inclusiveness, and alignment with open science norms.

Clarifying the Interdependence of Strategies

The five strategies outlined below are designed to operate as a coordinated system. While each focuses on a specific function, they are interdependent components of a shared infrastructure for evidence synthesis and reuse.

- Strategy 1 (Federated Repository) defines the **system-level goal**: an interoperable, distributed network of evidence repositories that supports structured access, contribution, and reuse.
- Strategy 2 (Record Structures and Formats) focuses on the **technical structuring** of evidence data—how information is formatted, encoded, and transformed for reuse across systems.
- Strategy 3 (Metadata Standards) addresses the **semantic and discovery layer**, enabling classification, filtering, and linkages across repositories using multilingual taxonomies, identifiers, and tagging tools.
- Strategy 4 (Open Access Standards) centres on the **legal and policy frameworks**—such as licensing, attribution, and permissions—that enable equitable sharing and reuse.
- Strategy 5 (Quality Assurance) functions **across all other strategies**, supporting consistent validation, trust, and system-wide coordination.

Together, these strategies enable interoperability across platforms and across the lifecycle of evidence—from structured extraction to discoverable metadata to open, reusable content. Where appropriate, ESIC will adopt or adapt external standards (e.g., FHIR, RDF, OAI-PMH, AGROVOC) and create specific profiles tailored to the needs of evidence synthesis. Overlaps are intentional to ensure integrated delivery, but implementation efforts will assign clear functional scopes to avoid redundancy.

Strategy 1: Building a Federated Repository of Living Evidence Data <i>Evidence synthesis infrastructure (tools, platforms)</i>	
Problem	<p>Despite the growing volume of evidence, its effective use is hindered by systemic fragmentation and access limitations.</p> <ul style="list-style-type: none"> Existing evidence repositories operate in silos, limiting cross-platform coordination. Many repositories restrict access—sometimes due to legal or institutional barriers that prevent data sharing across borders or in specific formats. There is widespread duplication of effort in data extraction and synthesis stages. The absence of a connected and interoperable systems prevents coordinated discovery and reuse
Description	<p>This strategy proposes a connected system of evidence repositories that operate in coordination, rather than through a single centralized platform. Each repository retains control over its own data while contributing to broader reusability by aligning, where feasible, with shared standards and interoperability practices. Discovery services may include optional aggregation of basic metadata or other mechanisms to support navigation across repositories. The aim is to make structured evidence easier to find, use, and build upon—while respecting diverse data policies and governance models.</p> <p>To support equitable participation, the strategy includes dedicated measures for institutions that lack the capacity to build or maintain their own repositories. These may include lightweight open-source tools, shared hosting options, and technical onboarding support. These supports lower entry barriers and enable LMIC institutions, smaller research groups, and non-traditional contributors to participate directly in the global evidence infrastructure..</p>
Expected outcomes	<ul style="list-style-type: none"> Lower infrastructure and maintenance costs for funders and institutions by aligning systems around shared standards, APIs, and governance models. Improved support for living evidence synthesis, as connected repositories and real-time updates enable ongoing integration of new findings into synthesis workflows. Reduced duplication of effort in evidence synthesis, as shared infrastructure increases visibility and reusability of existing data across systems. Increased participation by LMIC institutions in contributing to, accessing, and applying shared evidence, with measurable growth expected within the first two years. Broader access to underrepresented evidence for LMIC institutions, multilingual users, and regional decision-makers through regionally governed, multilingual nodes. Faster and more targeted discovery of evidence for reviewers, platforms, and policy users through structured search across interoperable repositories.
Costing	<p>Estimated at USD 26.4 million, with 71% allocated to staffing for mapping, curating, governing, and technically linking distributed repositories. 27% supports contracts for platform development, AI tool integration, and infrastructure testing. The remaining costs cover tools, travel, and stakeholder engagement (2%).</p> <p>The distribution reflects the system-level complexity and hands-on nature of building an interconnected, globally inclusive evidence infrastructure.</p>
Value for Money	<p>This strategy addresses a core barrier to evidence reuse: fragmented and inaccessible data. By enabling structured, cross-repository access, it reduces duplication of extraction and formatting work, which currently consumes significant time and resources.</p> <p>The investment supports shared infrastructure that lowers long-term costs for both well-resourced and under-resourced actors. It allows evidence synthesis teams to build on existing data rather than repeat foundational work—improving speed, efficiency, and consistency.</p> <p>Returns are system-wide: faster evidence availability for decision-makers, lower duplication for researchers, and reduced maintenance costs for funders and platforms. By connecting distributed repositories and enabling real-time updates, this strategy provides the technical backbone for living evidence synthesis—where evidence can be continuously updated, discovered, and reused at scale.</p>
Equity considerations	<p>This strategy supports equitable participation by allowing local, regional, or sector-specific repositories to operate. This strategy supports participation from institutions with different capacities by allowing repositories to maintain their own data policies while contributing to a shared infrastructure. Governance will emphasize shared leadership, with LMIC and community-based institutions actively involved in standard-setting processes.</p>

To reduce structural barriers, the strategy includes multilingual interfaces, open APIs, and lightweight tools. Offline access, caching, and localization workflows will also be explored during prototyping to support use in low-bandwidth environments.

By embedding equity into governance and infrastructure design, the strategy enables underrepresented institutions not only to access and contribute to global evidence, but also to take an active role in shaping how the system evolves.

Description of core components:

Convening a federated repository of living data involves building an interconnected system of repositories—each maintaining its own governance and structure, while participating in shared standards and infrastructure. The system will include discovery functions to help navigate and access evidence across repositories. The overall system will support interoperability, equity, and local control, while enabling shared access.

At least seven core components will be part of this initiative:

- **Mapped and Curated Repository Network:** A living map of evidence data repositories, organized by sector, coverage, evidence type, and question type. This includes identifying and curating existing repositories to surface gaps, reduce duplication, and support future linkages.
- **Federated Nodes for Sovereign or Restricted Data:** Interoperable regional or sectoral repositories for cases where data cannot be exported or centrally stored due to policy, legal, or ethical considerations. These nodes are linked to the central system while respecting local ownership.
- **Open API for Cross-Platform Data Integration:** A standardized, open-access application programming interface (API) to enable consistent data exchange across existing and emerging platforms, facilitating real-time integration and interoperability. This also allows tools like RevMan to connect directly, so users can access or be alerted to new evidence within their workflow—supporting living evidence processes.
- **AI-Enabled Infrastructure for Data Extraction and Curation:** Artificial intelligence tools to support the extraction, structuring, and curation of data entering the repository. These tools assist in preparing content for reuse and improving navigability across repositories.
- **User Interface for Data Exploration and Analysis:** An accessible, multilingual interface focused on helping users—especially non-technical audiences—search, filter, and access structured evidence easily across repositories. The interface prioritizes discoverability and reuse, rather than in-depth data analysis..
- **Multilingual Access and Navigation:** All repository functionalities will be accessible in multiple languages to reduce access barriers and ensure inclusive usability. This includes multilingual interfaces for search, filtering, with built-in machine translation to help users quickly assess the relevance of evidence in their own language.
- **Equity-focused infrastructure support:** Provision of lightweight open-source tools, shared hosting options, and onboarding support for institutions that lack the technical capacity to build their own repositories—ensuring equitable participation in the federated system.

Key Activities and Sequencing:

This strategy will be implemented through a phased approach.

- **Phase 1: Foundation & Prototyping** will include mapping existing repositories, establishing governance and technical working groups, compiling and validating existing Open APIs and developing additional APIs as needed, and compiling, validating, and adapting AI tagging tools and reuse interfaces. This initial phase will also involve identifying pilot LMIC institutions for early participation.
- **Phase 2: Expansion & Feature Enrichment** will focus on launching 2–3 federated nodes and enhancing existing repository connections. Additional activities will include expanding AI tools with feedback loops, building multilingual interfaces and initiating translation workflows, and developing user-driven data analysis tools with flexible export options for users to apply their own tools. This phase will also involve starting structured training via WG5.
- **Phase 3: Maturity & Integration** will aim to achieve full technical interoperability across nodes, institutionalize governance, metrics, and update cycles, operationalize translation infrastructure and workflows, integrate the repository with global platforms and policymaking tools, and launch a long-term funding and sustainability strategy.

Strategy 2: Developing Standardized Record Structures and Interoperable formats <i>Evidence synthesis infrastructure (tools, platforms)</i>	
Problem	<p>Interoperability across evidence synthesis systems is hindered by:</p> <ul style="list-style-type: none"> • Lack of consistent/standardized record structures for different inputs and outputs of evidence syntheses. • Incompatible or informal data formats, preventing reuse and platform integration • Vast amounts of legacy data locked in non-standard or narrative-only formats, inaccessible for automated processing.
Description	<p>This strategy addresses structural fragmentation by developing shared record structures and machine-readable formats for evidence synthesis data. It supports both prospective and retrospective conversion of data into standardized, interoperable formats.</p> <p>The focus is on how synthesis inputs (e.g., primary studies) and outputs (e.g., data extractions, appraisals) are organized and exchanged across systems, not on harmonizing source document formats. Some alignment with source reporting may be explored where feasible.</p> <p>The goal is to ensure both standardization (for consistency) and interoperability (for cross-platform compatibility), providing a structural foundation for broader reuse across organizations and sectors. These formats will support routine evidence synthesis tasks such as extracting structured data for meta-analysis, quality appraisal, or decision dashboards.</p>
Expected outcomes	<ul style="list-style-type: none"> • Faster integration of evidence into synthesis workflows and platforms, reducing manual formatting and enabling smoother, more automated processes across systems. • Lower technical barriers for participation by LMIC institutions and smaller research teams through open, standardized formats. • Improved consistency and transparency in how evidence is structured and exchanged, enabling provenance tracking and better integration with metadata systems. • Greater long-term interoperability across the evidence ecosystem, supporting platform developers, funders, and cross-sector users with a shared technical foundation. • Stronger system-level infrastructure for evidence synthesis, reducing duplication and enabling sustained reuse across projects, organizations, and contexts.
Costing	<p>Estimated at USD 5.26 million, with 90% allocated to staffing for developing shared formats, conversion tools, and technical protocols. The remaining costs support training (6%), travel (3%), and tooling.</p> <p>The distribution reflects the hands-on, iterative nature of technical standards development.</p>
Value for Money	<p>This strategy improves the speed, quality, and efficiency of evidence synthesis by reducing manual formatting and enabling seamless data transfer between systems. It supports both automation and reuse—lowering costs and duplication across projects.</p> <p>By enabling structured, machine-readable records, the strategy increases the return on existing data investments and reduces integration costs for institutions. Funders benefit from enhanced reuse, discoverability, and readiness for downstream applications, including AI-supported synthesis and decision tools.</p> <p>Support for both current and legacy data allows phased adoption and immediate use. Alignment with open standards reduces vendor lock-in and supports scale across geographies and systems.</p>
Equity considerations	<p>Standardized formats lower the technical barriers to entry, enabling LMIC institutions and smaller organizations to contribute and reuse data using open tools and templates. Retrospective conversion efforts ensure data produced in low-resource settings is not excluded from the shift toward structured and automated synthesis workflows.</p> <p>Supporting multilingual documentation and community co-design will further strengthen accessibility and equity.</p>

Description of core components:

At least four core components will be part of this initiative:

- A standardized record structure for evidence synthesis inputs (e.g., primary studies, evaluation reports) and outputs (e.g., synthesis reports, quality appraisals, data extracted, etc.).

- Promoting the use and alignment of interoperable machine-readable data formats by adapting existing standards to support free exchange across platforms in the evidence synthesis ecosystem.
- Systems to support the prospective conversion of newly generated data into the defined interoperable formats, helping standardization become part of routine workflows..
- Systems and tools to enable retrospective conversion of legacy data into the interoperable format, with an emphasis on high-priority datasets and scalable approaches that balance impact and feasibility
- Multilingual documentation and co-design resources will also be developed to support adoption, particularly by LMIC institutions and smaller organizations.

Key Activities and Sequencing:

This strategy will be implemented through a phased approach.

- In **Phase 1: Foundation & Prototyping**, activities will include establishing technical and advisory groups to oversee standard development, mapping existing standards and tools for evidence data sharing, developing initial record structure standards for inputs and outputs, drafting interoperable data formats and exchange protocols, and prototyping tooling for data conversion and structure validation.
- **Phase 2: Expansion & Feature Enrichment** will focus on building and testing retrospective conversion toolkits. This phase will also involve piloting conversion workflows with selected institutions, such as LMIC partners and sectoral repositories.
- Finally, **Phase 3: Maturity & Integration** will aim to launch open consultation and finalize standards based on feedback. It will also institutionalize standards through integration into platform governance and training programs.

Strategy 3: Establishing Metadata standards to facilitate data identification and discoverability <i>Evidence synthesis infrastructure (tools, platforms)</i>	
Problem	<p>Evidence remains difficult to find and reuse without standardized metadata, even when the underlying data is structured. Key issues include:</p> <ul style="list-style-type: none"> • Missing or inconsistent metadata, limiting discovery and interoperability • Lack of shared identifiers, classification systems, and descriptive tags across platforms • Over-reliance on time-consuming, error-prone free-text search by systematic reviewers • Limited ability to filter by topic, population, method, or outcome type • Difficulty tracking provenance, version history, and quality status—reducing trust • Under-indexing of grey literature, non-English sources, and LMIC-generated evidence
Description	<p>This strategy aims to develop and implement metadata standards that enable consistent classification, discoverability, and linkage of data used or produced in evidence syntheses. It will enhance interoperability across platforms by supporting structured tagging of studies, reports, and other outputs.</p> <p>This includes adopting or developing common typologies and taxonomies—organized by topic, method, population, and context—to support multilingual filtering, classification, and reuse across federated platforms. Metadata will be integrated upstream into workflows to support automation, provenance tracking, and cross-platform interoperability.</p>
Expected outcomes	<ul style="list-style-type: none"> • Stronger transparency, trust, and accountability across the system, enabled by metadata fields that track provenance, version history, and curation steps. These features allow users to assess the trustworthiness and evolution of evidence across platforms. • Reduced workload and workflow bottlenecks for evidence producers, research teams, and platform developers, as metadata standards support upstream tagging, automation, and consistent reuse. • Faster and more accurate evidence retrieval for systematic reviewers, data scientists, and decision-makers, enabled by precise filtering by topic, population, method, and outcome. • Improved visibility and reuse of underrepresented evidence from LMIC institutions, multilingual users, and civil society organizations, through multilingual tagging and standardized classifications applied to grey literature and non-indexed outputs.

Costing
Total cost is USD 6.37 million, with 82% allocated to staffing for schema development, multilingual tagging, and classification systems. Most of the remaining budget supports training (15%) and pilot testing. This allocation reflects the collaborative and cross-sector effort needed for inclusive metadata design.
Value for Money
<p>Metadata standards deliver strong returns by making evidence easier to find, classify, and reuse—without changing the content itself. They reduce screening time, support automation, and improve integration of diverse sources, including grey literature and LMIC outputs.</p> <p>By enabling structured tagging and version tracking, this strategy supports living evidence synthesis—allowing updates to be efficiently discovered and integrated. The approach is low-cost, scalable, and compatible with existing systems.</p> <p>Shared taxonomies, multilingual vocabularies, and persistent identifiers further enhance discoverability, equity, and cross-platform interoperability with minimal ongoing cost.</p>
Equity considerations
Metadata standards will be inclusive by design, with multilingual fields and community input. By increasing the visibility of underrepresented outputs—such as non-English literature, LMIC evaluations, and grey literature—this strategy supports equitable access to evidence. It also lowers technical barriers to participation by providing templates and tagging tools suitable for institutions with limited infrastructure.

Description of core components:

The following components are essential to ensure consistent, multilingual, and interoperable metadata practices across platforms, enabling effective discovery, filtering, and reuse of evidence:

- Modular metadata Schema and Classification Standards:** A metadata schema covering study/article IDs, classifications (e.g., design, population, outcome type), thematic tags (e.g., SDGs), and quality fields. The metadata should be modular, allowing to be adapted from different types of records (inputs and outputs of evidence syntheses) and platforms. To reduce bottlenecks for evidence producers, the schema will prioritize essential fields and support incremental enrichment over time. Tagging should be as close to upstream workflows as possible and supported by automation where feasible. Where feasible, a common core module can be established to promote cross-sector interoperability, while allowing additional sector-specific or regional modules to extend the schema based on local needs or evidence types. This balance supports consistency without imposing uniformity.
- Multilingual and Cross-Sector Alignment of typologies:** Controlled vocabularies and metadata fields translated and validated across multiple languages and sectors (e.g., health, education, environment). This alignment enables consistent classification, filtering, and discovery of evidence across domains, while ensuring accessibility for diverse users. Where appropriate, sector-specific taxonomies should be mapped to shared typologies to support cross-sector interoperability without sacrificing domain relevance.
- Tools that enable transformation to support taxonomy alignment:** Manual and AI-assisted tools for applying metadata, with built-in QA features to ensure tagging is consistent, traceable, and updateable, and can be translated to multiple available taxonomies. These tools should be designed to minimize burden on producers and facilitate reuse of metadata across platforms, including alignment with external standards (e.g., Dublin Core, Schema.org) to reduce duplication.
- Transparent Provenance and Versioning Records:** Metadata fields that record changes, updates, and curation steps—allowing users to understand what transformations the data has undergone.
- Identifier Systems and Linkages:** Guidance and tooling for assigning and maintaining persistent identifiers (e.g., DOIs), with linkages between studies, syntheses, and derivative outputs.
- Metadata Enrichment of Legacy Records:** Where feasible, tagging standards may be applied to previously structured records to enhance discoverability.

Key Activities and Sequencing:

This strategy will be implemented through a phased approach.

- In **Phase 1: Foundation & Prototyping**, activities will include establishing a metadata working group and regional advisors, defining the metadata schema and classification system, mapping and reviewing existing metadata standards and typologies, and developing a strategy for mapping legacy metadata into a unified framework. This phase will also involve drafting identifier guidance and linkage protocols.
- **Phase 2: Expansion & Feature Enrichment** will focus on developing tagging tools, including AI-assisted modules. This phase will also involve conducting multilingual validation and testing, and piloting the tagging system across selected ESIC-aligned repositories.
- Finally, **Phase 3: Maturity & Integration** will aim to finalize standards through open consultation and feedback. It will also institutionalize the metadata schema and QA practices across ESIC tools, and maintain the multilingual metadata dictionary and update protocols.

Strategy 4: Advancing Open access standards for equitable data sharing and reuse <i>Evidence synthesis infrastructure (tools, platforms)</i>	
Problem	<p>Despite increasing volumes of evidence and synthesis data, access remains restricted due to legal, technical, and policy barriers. Without shared standards, open access practices remain fragmented and inequitable. Key challenges include:</p> <ul style="list-style-type: none"> • Unclear or restrictive licensing terms that limit reuse and redistribution • Paywalls or subscription models that create access inequities • Lack of consistent norms for attribution, versioning, and derivative use • Fragmented policies across repositories, institutions, and publishers • Limited infrastructure to support equitable open access across languages and sectors
Description	<ul style="list-style-type: none"> • This strategy aims to advance open access standards and policies to enable equitable sharing and reuse of evidence synthesis data across geographies, sectors, and user groups. It addresses legal and institutional barriers by clarifying data ownership, licensing, and ethical reuse frameworks. • While the primary focus is integrating these practices within ESIC platforms and governance structures, the strategy also contributes to global open science dialogues—positioning evidence synthesis as a critical use case for equitable knowledge sharing. Rather than duplicating mandates, it adapts existing standards and platforms to serve diverse contexts and users more inclusively.
Description	<p>This strategy aims to advance open access standards and policies to promote the equitable sharing and reuse of data that supports evidence-based decision-making across diverse contexts, sectors, and user groups. It addresses legal, structural, and institutional barriers that limit data use, clarifies data ownership and licensing arrangements, and encourages fair practices that benefit everyone globally. While the primary focus is on embedding these standards within the ESIC ecosystem—through platform governance mechanisms and institutional practices—the strategy also contributes to global policy influence by aligning with and supporting ongoing international initiatives in the open science and access community. Rather than duplicating mandates, it seeks to elevate the unique needs of evidence synthesis as a use case for equitable knowledge sharing.</p>
Expected outcomes	<ul style="list-style-type: none"> • Alignment with open science and funder mandates, enabling institutions to comply with access requirements while promoting public-good knowledge sharing. • Scalable open access models that reduce paywalls and permission barriers—particularly benefiting early-career researchers, LMIC platforms, and local evaluators. • Increased participation from diverse institutions, especially LMIC-based and under-resourced actors, in accessing and sharing evidence. • Broader and more equitable reuse of evidence, as standardized licensing reduces legal and technical barriers across geographies and sectors. • Greater transparency and accountability, with standard attribution and versioning improving visibility into evidence origin and downstream use. • Global recognition of ESIC’s open access approach as a replicable model through engagement with international actors and alignment with evolving best practices.

Costing
Total cost is USD 4.22 million, with 65% allocated to staffing and 23% to legal and technical contracts for licensing and access frameworks. The remainder supports training, travel, and outreach. This reflects the policy and coordination focus required to embed open access in platforms.
Value for Money
<p>This approach maximizes returns on public investment by embedding clear, standardized licensing and ethical reuse practices into ESIC platforms. It reduces legal ambiguity, avoids duplication, and removes the need for fragmented, institution-specific access policies, lowering long-term legal, administrative, and compliance costs.</p> <p>By aligning access protocols with global open science initiatives and integrating them into platform governance, the strategy strengthens the infrastructure for scalable, equitable data sharing.</p> <p>It also enhances the foundations for living evidence synthesis by enabling faster incorporation of new findings, broader redistribution, and more inclusive participation particularly by LMIC actors and multilingual users. These efficiencies collectively support a more connected, cost-effective, and sustainable evidence ecosystem.</p>
Equity considerations
<p>Open access is a key enabler of equity, removing financial, legal, and technical barriers that often prevent institutions and individuals—particularly in LMICs—from participating fully in the global evidence ecosystem. This strategy embeds licensing clarity, ethical reuse practices, and access protocols into ESIC systems to ensure that publicly funded evidence can be reused by all, regardless of geography, institutional size, or language.</p> <p>Multilingual access, alignment with national data rules, and support for responsible local ownership are core to the approach. Inclusive governance structures will ensure that LMIC and community-based repositories have a voice in how open access standards evolve. In doing so, the strategy recognizes and amplifies successful models already developed in under-resourced contexts, positioning them as contributors to global best practice. By reducing dependence on informal or pirated access channels, this strategy promotes legal, ethical, and equitable access to high-quality evidence for all.</p>

Description of core components:

The following components define the building blocks of an equitable open access ecosystem for data sharing and reuse:

- **Open Licensing Standards:** Guidance on recommended licenses (e.g., CC-BY, CC0) and conditions of reuse and text data mining, adapted for synthesis data and evidence workflows.
- **Infrastructure Support for Access:** Platform-level capabilities for downloading, linking, versioning, and translating open datasets across languages and formats, and supporting organizations in making all their products open access.
- **Legal and Ethical Guidance for Reuse:** Templates and policies to ensure compliance with privacy, IP, and data protection laws—particularly when sharing sensitive or context-specific evidence.
- **Open Access Integration into Platform Governance:** Roles and responsibilities for implementing, monitoring, and updating open access practices, embedded in repository and WG-level oversight structures. Oversight will include mechanisms for periodic review, feedback integration, and alignment with evolving institutional policies, particularly in LMIC contexts.

Key Activities and Sequencing:

This strategy will be implemented through a phased approach.

- In **Phase 1: Foundation & Prototyping**, activities will include convening an open access advisory group, drafting open licensing and attribution guidance, and conducting a legal review of reuse models across jurisdictions.
- **Phase 2: Expansion & Feature Enrichment** will focus on building platform capabilities for license display, citation export, and open access tagging. This phase will also involve developing and piloting multilingual access protocols, and hosting consultations with LMIC institutions and data producers on access priorities.

- Finally, **Phase 3: Maturity & Integration** will aim to integrate licensing and attribution norms into platform governance frameworks, monitor and evaluate open access implementation across ESIC systems, and maintain public guidance materials and licensing templates.

Strategy 5: Ensuring Quality and Monitoring of Data Sharing and Reuse Systems Evidence synthesis process (methods; training; learning; sharing; convening)	
Problem	<p>Data sharing systems often lack built-in mechanisms to assess and maintain the quality, relevance, and ethical use of the data they host. Without systematic M&E, users face difficulties in judging whether shared data is current, accurate, or responsibly curated.</p> <p>Key challenges include:</p> <ul style="list-style-type: none"> • Lack of common QA processes for evaluating shared data • Inconsistent application of quality standards across repositories • Absence of feedback loops for reporting and correcting data issues • Limited ability to monitor system usage, equity, and performance • Unclear accountability for data quality, ethics, and legal compliance
Description	<p>This strategy aims to embed a tiered validation system alongside quality assurance (QA), monitoring, and evaluation (M&E) mechanisms within the infrastructure and processes that support data sharing and reuse in evidence synthesis. The tiered approach enables different levels of review, ranging from automated checks and methodological audits to participatory peer review, matched to the risk profile, context, and intended reuse of the data. It seeks to ensure that shared data is not only accessible but also trustworthy, up to date, ethically managed, and reusable with confidence by evidence producers and users.</p> <p>Recognizing that quality assurance and evaluation processes will never meet all expectations in all contexts, the strategy also emphasizes continuous improvement. It establishes mechanisms to assess system performance, equity impacts, and adherence to shared standards, while leaving room to resolve tensions and refine evaluation over time.</p>
Expected outcomes	<ul style="list-style-type: none"> • Improved system-wide tracking and accountability for funders, governance bodies, and coordination platforms, with M&E frameworks generating actionable insights on usage, equity, and adherence to shared standards. • More efficient reuse of high-quality evidence for reviewers, curators, and guideline developers, as tiered QA systems help filter by rigor, relevance, and readiness. • Faster identification and correction of data issues for contributors, repository managers, and QA teams, supported by feedback loops and user reporting mechanisms. • Increased trust and confidence in shared data among decision-makers, synthesis platforms, and evidence users, driven by transparent validation processes and visible quality indicators.
Costing	<p>Total cost is USD 13.72 million, with 97% allocated to staffing for QA systems, participatory review, and monitoring infrastructure. Tooling and travel represent a very small share. The cost profile reflects the intensive expert input needed to ensure system-wide trust and quality.</p>
Value for Money	<p>Embedding quality assurance and monitoring mechanisms directly into infrastructure strengthens the overall evidence ecosystem. It helps avoid costly duplication, data errors, and loss of trust—ensuring that shared data remains usable, up to date, and ethically sound across platforms and contexts.</p> <p>This strategy enhances the efficiency and reliability of living evidence synthesis by enabling users to rapidly identify high-quality, curated evidence and track updates over time. Tiered validation systems, feedback loops, and participatory review processes reduce the need for repeated verification, while supporting continuous learning and refinement.</p> <p>By integrating performance tracking and equity monitoring into platform governance, the approach delivers system-level returns—supporting funder accountability, improving user confidence, and enabling better decision-making across diverse regions and sectors..</p>
Equity considerations	

Robust QA and M&E systems help ensure that data from LMICs is not unfairly flagged or excluded due to differences in format or reporting practices. By developing transparent, locally informed standards, and involving LMIC actors in defining quality benchmarks and M&E indicators, this strategy builds inclusive trust. It also supports monitoring equity in access, participation, and benefit-sharing across data systems.

Expected outcomes

- Improved trust and reliability of shared data for decision-makers, synthesis platforms, and evidence users, through transparent validation processes and visible quality indicators.
- More efficient reuse of high-quality evidence for reviewers, curators, and guideline developers, as tiered QA systems help users filter by relevance, rigor, and readiness for decision-making.
- Faster identification and correction of quality issues for data contributors, repository managers, and QA teams, enabled by integrated feedback loops and user-driven reporting mechanisms.
- Better system-wide performance tracking for funders, governance bodies, and coordination platforms, with M&E frameworks generating data on usage, equity, and adherence to shared standards.

Description of core components:

The following components ensure that systems for data sharing and reuse remain high-quality, trusted, and responsive to user needs:

- **Tired validation Standards and Guidelines:** Co-developed criteria and templates to assess the completeness, relevance, reliability, and ethical compliance of shared data.
- **Monitoring and Evaluation Framework:** A set of indicators, dashboards, and review mechanisms designed to assess data usage, validation uptake, access equity, and alignment with shared standards. This includes tools to gather structured feedback from users and contributors, mechanisms to surface emerging issues, and processes for governance bodies to act on evaluation findings. Recognizing the diversity of platforms and contexts, the framework will be implemented with flexibility, allowing for local adaptation and periodic refinement.
- **Feedback and Reporting Channels:** Tools to allow users to flag issues with data quality, usage barriers, or policy misalignment—feeding into continuous improvement processes.
- **Participatory Review Processes:** Structured opportunities for interest-holder input on what quality and ethical data use means across sectors and regions.
- **Integration with System Governance:** Clear responsibilities for implementing QA/M&E protocols at both the platform and institutional levels, aligned with WG2-wide governance models.

Key Activities and Sequencing:

This strategy will be implemented through a phased approach.

- In **Phase 1: Foundation & Prototyping**, activities will include establishing a technical QA and M&E advisory group, developing draft quality assurance standards, and defining indicators for system monitoring and ethical adherence.
- **Phase 2: Expansion & Feature Enrichment** will focus on building user feedback and issue-reporting interfaces. This phase will also involve piloting QA and M&E systems in selected repositories, and conducting participatory review sessions with LMIC and sectoral partners.
- Finally, **Phase 3: Maturity & Integration** will aim to integrate QA/M&E protocols into platform governance and user workflows, maintain and update monitoring dashboards and QA tools, and conduct periodic system-wide quality and equity evaluations.

CROSS-CUTTING STRATEGY

In addition to the five technical strategies, the report identifies a critical cross-cutting condition: Sustainable Funding and Incentive Alignment. This strategy is foundational to the long-term success, equity, and scalability of the entire infrastructure. It is considered “cross-cutting” because its impact is not limited to a single domain—instead, it underpins the implementation, sustained use, and institutional adoption of all other strategies. Without reliable financing and well-aligned incentives, even the most innovative efforts—whether federated repositories, standardized formats, enriched metadata, or open access systems—are unlikely to achieve widespread uptake, maintenance, or equitable participation, particularly from LMICs and under-recognized contributors.

Sustaining the Infrastructure: Funding Models and Incentives <i>Evidence synthesis infrastructure (tools, platforms)</i>	
Problem	<p>Despite broad recognition of the value of evidence sharing and synthesis, funding remains fragmented, short-term, and project-based. This undermines the scalability and sustainability of infrastructure investments.</p> <p>Key challenges include:</p> <ul style="list-style-type: none"> • Lack of long-term funding to maintain and evolve core infrastructure • Fragmentation of funding streams, leading to duplication and inefficiencies • Absence of pooled or joint mechanisms to support shared infrastructure • Limited incentives for data contributors, curators, or LMIC institutions • Mismatched reward systems that prioritize publication over sharing and reuse <p>Even where infrastructure exists, uptake is often hindered by low motivation, institutional inertia, and cultural resistance.</p>
Description	<p>This cross-cutting strategy focuses on establishing the long-term financial and motivational systems needed to sustain the infrastructure, processes, and collaborations envisioned in the ESIC planning process. It aims to secure sustainable, multi-source funding and to align incentives across sectors, regions, and stakeholder groups to ensure ongoing participation in data sharing, reuse, and system upkeep.</p> <p>In addition to financial mechanisms, this strategy addresses the behavioral and cultural dimensions of infrastructure use—recognizing that incentives are a key lever for changing norms and encouraging consistent participation across context.</p>
Expected outcomes	<ul style="list-style-type: none"> • Multi-donor pooled funds and shared grant frameworks supporting infrastructure over time. • Transparent cost models and sustainability plans embedded in each platform. • Reward systems that recognize and incentivize data sharing and reuse. • Equitable access to funding opportunities across geographies and institution types. • Reduced duplication and greater coordination among funders and stakeholders <p>Strengthened voluntary institutional alignment through coordinated advocacy and stakeholder engagement, supporting the uptake of shared standards, tools, and governance models.</p>
Costing	<p>This cross-cutting strategy does not have a standalone budget line. Instead, its implementation relies on components embedded within other strategies, including shared funding mechanisms, platform-specific sustainability planning, and pilot incentive schemes.</p> <p>Related costs are distributed across the staffing, governance, and engagement activities in Strategies 1–5.</p>
Value for Money	<p>Sustainable infrastructure requires sustainable investment. This strategy delivers strong long-term value by establishing pooled funding models, reducing fragmentation, and aligning donor efforts to support global public-good infrastructure. Clear cost models and coordinated funding reduce duplication, improve transparency, and enable platforms to evolve without restarting from scratch each funding cycle.</p> <p>By embedding incentives into research assessment, funding decisions, and platform performance metrics, this strategy also increases returns on investment in human capital—recognizing and rewarding curators, translators, and LMIC institutions who are essential for data stewardship but often overlooked. Behavioral incentives further extend value by promoting actions that improve reusability and system integrity at low additional cost.</p> <p>Together, these approaches make infrastructure investments more resilient, equitable, and impactful—ensuring that participation is not just possible but sustainable and rewarding across all regions and roles.</p>
Equity considerations	<p>Equity is central to both the funding and incentive models. This strategy emphasizes funding models that actively include LMIC partners—not only as beneficiaries, but as co-leaders in governance, implementation, and innovation. Incentive systems will also reflect contributions from curators, translators, local reviewers, and regional institutions that are often under-recognized but essential for system functioning</p>

Description of Core Components:

The following components support the long-term sustainability and motivation needed to make the ESIC infrastructure viable and inclusive:

- **Shared Funding Mechanisms and Sustainability Models:** Multi-donor pooled funds, aligned investment structures, and platform-specific business models—including cost-sharing and public-good justifications—to support long-term infrastructure, governance, and operations.
- **Aligned Incentives and Equitable Access:** Recognition and reward systems for contributors (e.g., data producers, translators, curators), integrated into funding and assessment frameworks. Includes mechanisms to ensure equitable access to grants, leadership roles, and capacity support for LMIC institutions and early-career professionals.
- **Coordination and Accountability Structures:** Mechanisms to align funders and stakeholders, reduce duplication, and ensure transparent monitoring of progress, sustainability, and impact across platforms and partners.
- **Behavioural Incentive Models and Culture Change Tools:** Tools and policies that motivate alignment with shared infrastructure—such as tagging structured data, adopting core outcome sets, and using standard formats. These promote long-term engagement and system stewardship.
- **Advocacy and Engagement Functions:** Cross-cutting activities to build institutional trust, support voluntary alignment, and drive uptake of shared practices. Includes coordinated messaging, stakeholder partnerships, and co-creation approaches across regions and sectors.

Key Activities and Sequencing:

This cross-cutting strategy will be implemented through a phased approach.

- In **Phase 1: Foundation & Prototyping**, activities will include mapping existing funding flows and gaps across ESIC-aligned infrastructure, convening funders to design pooled or coordinated funding models, and drafting an incentive framework and potential recognition models.
- **Phase 2: Expansion & Feature Enrichment** will focus on piloting shared grant mechanisms across multiple funders. This phase will also involve developing costing templates and sustainability plans for each platform, launching incentive pilots (e.g., citation tracking, recognition in grant calls), and co-designing behavioural incentive pilots (e.g., recognition for data curation, open tagging, or use of core outcome sets) with institutions and funders.
- Finally, **Phase 3: Maturity & Integration** will aim to institutionalize funding and incentive models across ESIC platforms, monitor and evaluate sustainability and equity outcomes, and adapt and scale up funding and incentive models based on feedback.

CONCLUSION AND NEXT STEPS

This Stage 4b report presents five prioritized strategies from ESIC Working Group 2 to strengthen global systems for data sharing and reuse. Together, they offer a coordinated, equitable, and scalable foundation for interoperable, trusted evidence synthesis.

Costing estimates are now complete. Several strategies—especially on metadata, record structures, open access, and QA—are feasible within 24–36 months, with pilots possible in year one. More complex elements, like the federated repository, require a sustained effort but can begin within this timeframe.

The five WG2 strategies are estimated at USD 55.98 million, comprising the majority of the USD 56.2 million total projected for ESIC to date.

Next steps include open consultation and Stage 5 integration, culminating in the June 2025 consensus meeting where final proposals will be agreed.

Appendix 1. Costing

ESIC Total

Group	# of solutions	Group total costs (US\$)	Employment costs by location**					
			% Group person months			% Group total employment costs		
			GS -LLM	GH - HUM	GN	GS -LLM	GH - HUM	GN
PG1*	1	\$9,123,546	0%	100%	0%	0%	100%	0%
WG1*	5	\$72,386,225	7%	61%	32%	3%	52%	45%
WG2	5	\$55,979,378	34%	36%	29%	18%	33%	49%
WG3	7	\$33,956,375	33%	41%	25%	17%	42%	41%
WG4	11	\$42,210,240	35%	34%	31%	17%	33%	51%
WG5	10	\$40,200,646	47%	30%	23%	24%	32%	44%
ESIC total	39	\$253,856,410	26%	50%	24%	13%	49%	38%
Grand total with 5% M&E		\$266,549,230						
Range***	\$245,005,965	\$288,255,278						

*Includes just the preferred options. There are other less resource intensive options costed as well for PG1 and WG1.

**Average employment cost data was sourced from a multi-country analysis and other publicly available data, used in base case scenario.

***Considers lower and higher annual average **employment costs** than the ones used in base case; Low values (-19% in GS-LLM, -17% in GS-HUM and -3% in GN). Higher values 14%, 17% and 5% respectively, calculated using the full spectrum of available data.

Name of the group:	WG 2 Data Sharing and Reuse
--------------------	-----------------------------

ESIC Total incl M&E	\$266,549,230
Group Total	\$55,979,378

Range	
\$50,984,484	\$60,842,065

Option A: Total Group Employment costs by location

Location	Person months	% of person months	Employment costs	% of employment costs
Global South LLM	172	34%	\$6,669,189	18%
Global South HUM	183	36%	\$12,361,803	33%
Global North	147	29%	\$18,257,751	49%
Total	502	100%	\$37,288,743	100%

Solution Number	Total costs	% of Group Total
WG2.1	\$26,411,752	47%
WG2.2	\$5,262,661	9%
WG2.3	\$6,368,449	11%
WG2.4	\$4,221,489	8%
WG2.5	\$13,715,026	25%
Group Total	\$55,979,378	100%

Name of the group:	WG 2 Data Sharing and Reuse
Solution Number:	1
Solution Title:	Building a Federated Repository of Living Evidence Data

Federated LED Repository

Category	Direct Costs by category	Percentage of total direct costs
People	\$15,625,621	70.99%
Contracts	\$6,050,000	27.49%
Tools	\$165,900	0.75%
Training event attendance	\$35,833	0.16%
Citizen engagement	\$939	0.00%
Travel	\$115,300	0.52%
Dissemination	\$16,200	0.07%
Other	\$0	0.00%
Total direct costs	\$22,009,793	100.00%
Grand total with 20% overheads	\$26,411,752	

Name of the group:	WG 2 Data Sharing and Reuse
Solution Number:	2
Solution Title:	Developing Standardized Record Structures and Interoperable Formats

Standardised Structures

Category	Direct Costs by category	Percentage of total direct costs
People	\$3,965,679	90.43%
Contracts	\$0	0.00%
Tools	\$26,800	0.61%
Training event attendance	\$250,833	5.72%
Citizen engagement	\$939	0.02%
Travel	\$126,500	2.88%
Dissemination	\$14,800	0.34%
Other	\$0	0.00%
Total direct costs	\$4,385,551	100.00%
Grand total with 20% overheads	\$5,262,661	

Name of the group:	WG 2 Data Sharing and Reuse
Solution Number:	3
Solution Title:	Establishing Metadata standards to facilitate data identification and discoverability

Metadata standards

Category	Direct Costs by category	Percentage of total direct costs
People	\$4,344,259	81.86%
Contracts	\$45,000	0.85%
Tools	\$31,800	0.60%
Training event attendance	\$784,583	14.78%
Citizen engagement	\$1,899	0.04%
Travel	\$99,500	1.87%
Dissemination	\$0	0.00%
Other	\$0	0.00%
Total direct costs	\$5,307,041	100.00%
Grand total with 20% overheads	\$6,368,449	

Name of the group:	WG 2 Data Sharing and Reuse
Solution Number:	4
Solution Title:	Setting Open Access Standards for Equitable Data Sharing and Reuse

Open Access Standards

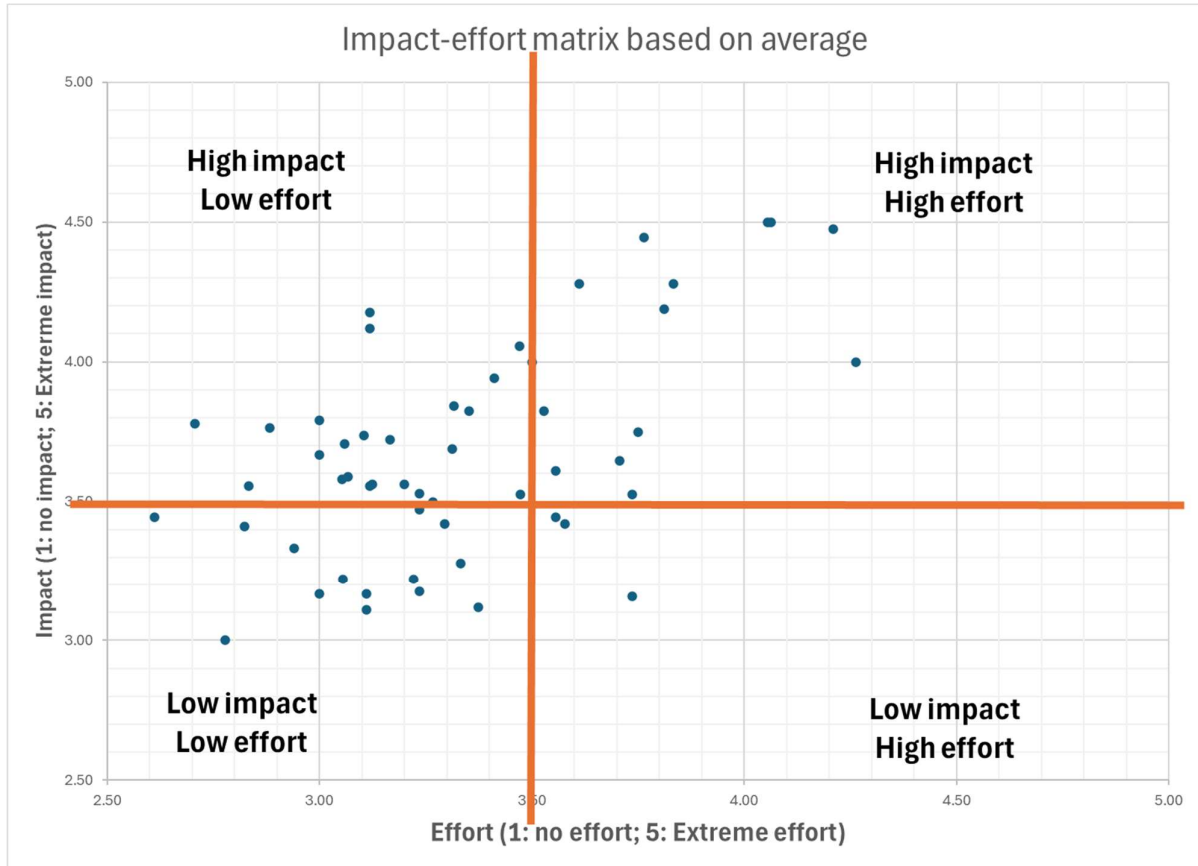
Category	Direct Costs by category	Percentage of total direct costs
People	\$2,302,035	65.44%
Contracts	\$820,000	23.31%
Tools	\$14,400	0.41%
Training event attendance	\$250,833	7.13%
Citizen engagement	\$939	0.03%
Travel	\$122,900	3.49%
Dissemination	\$6,800	0.19%
Other	\$0	0.00%
Total direct costs	\$3,517,908	100.00%
Grand total with 20% overheads	\$4,221,489	

Name of the group:	WG 2 Data Sharing and Reuse
Solution Number:	5
Solution Title:	Ensuring Quality and Monitoring of Data Sharing and Reuse Systems

Quality and monitoring

Category	Direct Costs by category	Percentage of total direct costs
People	\$11,051,149	96.69%
Contracts	\$170,000	1.49%
Tools	\$48,600	0.43%
Training event attendance	\$43,000	0.38%
Citizen engagement	\$939	0.01%
Travel	\$101,500	0.89%
Dissemination	\$14,000	0.12%
Other	\$0	0.00%
Total direct costs	\$11,429,188	100.00%
Grand total with 20% overheads	\$13,715,026	

Appendix 3. Impact-effort matrix



Appendix 4. Connection of strategies included in this report with the strategies included in report #3

Strategy	Strategies included from report #3
1. Building a federated repository of living evidence data	3.2.5 Deploy AI-enabled infrastructure to support tagging, content structuring, and extraction
	3.2.1 Federated repository of living evidence data
	3.2.2. Single living repository of data
	3.3.6 User-Friendly Tools for Non-Technical Data Reuse
	3.2.9. Creating comprehensive registries of protocols (of primary studies and evidence syntheses)
	3.2.4 Open API for data integration
	3.3.1 Multilingual user experience and access
	3.6.4 Strategic Funding for Foundational Dataset Production
2. Developing standardized record structures and interoperable formats	3.1.8 Develop standardized record structures for evidence synthesis inputs and outputs
	3.1.5. Evidence synthesis reporting standards
	3.1.6 Evaluation reporting standards
	3.1.7. Primary study reporting standards
	3.1.2 Interoperable data formats and exchange protocols
	3.1.9 Retrospective conversion of data and metadata
3. Establishing metadata standards for discovery and reuse	3.1.3 Global identifier framework
	3.1.1 Modular metadata standards (to facilitate data curation to be reused)
	3.1.4 Metadata transformation tools to support taxonomy alignment.
4. Advancing open access standards for equitable data sharing	3.3.4 Encouraging open licenses and publisher incentives for access
	3.3.5 Establishing organizational open access policies
	3.1.10 Open Access Standards for Data Sharing
	3.2.8 Supporting Open Access Infrastructure
	3.4.2 Encouraging open access through collaborative partnerships to share infrastructure and standard alignment
	3.7.4 Resources Supporting Organizational Shifts to Open Sharing (e.g., toolkits, pilots projects that show the benefits of open sharing)
	3.4.5 Advocacy for Text and Data Mining (TDM) Rights regulation
5. Ensuring quality and monitoring of data sharing systems	3.4.7 Monitoring and Evaluation of Data Sharing System Performance
	3.5.1 Tiered validation processes (to ensure quality and reliability of ES data)
	3.5.2 Methodological teams ensuring quality and replicability
Sustaining the Infrastructure Funding models and incentives	3.6.1 Multi-donor pooled funds to support ES infrastructure
	3.6.5 Sustainable funding for maintaining and updating living data infrastructure
	3.6.3 Recognition mechanisms to incentivize infrastructure contributions
	3.6.7 Link grant eligibility to data standards adherence
	3.6.9 Linked funding models for evidence lifecycle (incl funding for primary research)
	3.6.6 Equitable cost-sharing arrangements (incl models where different partners contribute in different forms)
	3.6.8 Funder and publisher incentives for data sharing
	3.4.6 Align Funding Models with Open Access Principles

Appendix 5. Strategies that were finally not prioritized as part of this report (note that some strategies were not necessarily excluded but reframed as part of other strategies included in this report)

3.1.11 Streamlined submission standards (single input, multiple uses)

3.1.5 Evidence synthesis reporting standards

3.2.3. Diagnostic pilots for data integration barriers

3.2.6 Provenance tracking systems

3.2.7 Technical deployment toolkits for platforms seeking integration.

3.3.2 Support offline access and low-bandwidth deployment

3.3.3 Regional mentorship models for co-development of platform strategies and synthesis infrastructure governance.

3.3.8. Creating translated version of ES outputs to facilitate reuses

3.4.1 Data-sharing and reusing agreement templates to build mutual trust

3.4.2 Collaborative partnerships for shared infrastructure governance and standard alignment

3.4.3 Incentives for conducting QA before data is shared

3.4.4 Align reuse systems with ethical and data sovereignty frameworks

3.5.3 Crowdsourced and peer-led quality assurance

3.5.4 Network of trust for institutional quality signaling

3.5.5 Re-use tracking tools using persistent IDs and contributor metadata

3.6.2 Coordinated funder platforms for infrastructure alignment

3.7.1 Regionally adapted training modules

3.7.2 Communities of practice for continuous learning and methodological alignment

3.7.3 Pilot demonstrations of reuse value

3.7.5 Localized data infrastructure hubs (that provide infrastructure support in LMI settings)

3.5.3 Crowdsourced and peer-led quality assurance

3.5.4 Network of trust for institutional quality signaling

3.5.5 Re-use tracking tools using persistent IDs and contributor metadata

3.7.2 Communities of practice for continuous learning and methodological alignment

3. Title of the solution:

Building a Federated Repository of Living Evidence Data

Activity	Role	Type	Location	Expected Number of staff	For how many months?	% Approx FTE
Mapping existing repositories	Research Analyst	Mid-tier		2	6	100
Manual curation of repository content	Content Curator	Mid-tier		9	60	100
Establish and Maintain governance and technical working groups	Governance Coordinator	Senior		1	60	100
				3		
Compile / validate existing Open APIs and develop additional ones, Build initial Open API	Software Developer / Data Architects	Senior			60	100
Develop AI tagging and reuse interface prototypes & expansion	Data Scientist / Engineer	Senior		3		
AI Applications	Product Manager	Senior		1	60	100
Designer (AI Applications)	Designer	Mid-tier		1	60	100
Identify pilot institutions to be apart of Federation	Engagement Lead	Mid-tier		1		
Launch federated nodes	product manager	Senior		3	60	100
Launch federated nodes	team lead	Senior		3	60	100
Launch federated nodes	senior back end engineers,	Senior		6	60	100
Launch federated nodes	senior infra engineers	Senior		6	60	100
	Data Infrastructure Specialist / System Integrators / Data Engineer			3		
Launch federated nodes		Senior			60	100
Senior Evidence/Domain/Regional Experts (Synthesis side judgement)	Evidence Expert	Senior		4		
Project management	Project Manager	Mid-tier		3	60	100
Federation Software development team	Front End Engineers	Senior		2	60	100
Federation Software development team	Full Stack Engineers	Senior		2	60	100
Federation Software development team	Data Engineers	Senior		2	60	100
Federation Software development team	Infrastructure Engineers	Senior		1	60	100
Federation Software development team	Product Manager	Senior		1	60	100
	Administrator/Coordinator	Junior		1	60	
Advisory Group	Advisors	Senior		10	60	10
Programme Management (Co ordinating between nodes)	Sr. Programme Manager			1	60	100
Evaluation and learning	Evaluation and learning expert	Senior		1	60	10
Launch long-term funding and sustainability strategy	Fundraising specialist	Senior		1	48	

Activity	Type of external service	Expected value of the contract (in US\$):	Delivering your solution, i.e. costs charged by an external supplier to deliver a piece of work? If so, please specify
Hosting cost for moderate-traffic node	Cloud hosting (yearly)	3000000	\$200K per repository per year (hosting)
AI training, validation, and/or inference			To Be deleted
	Translation services	500,000 (100,000 per year)	translation services/consultants should be \$100k/year at least
Translation services / Consultants		4000000	\$200K per repository per year (Maintenance staff & operational expenses) --> \$265K per repository per year (Maintenance staff & operational expenses)
Operations / Maintenance			

[illegible]

7. CAPACITY DEVELOPMENT: Are you proposing the solution includes funding to pay people (staff or others) to attend capacity building activities (e.g. short course/ training/CPD etc)?

Activity	Location of capacity building activity	Over the length of the solution, expected # of attendees funded (staff or others as part of the solution) to participate in the capacity building events	Average cost per attendee in US\$ (including fees, accommodation, travel, etc.) if known	OTHER REMARKS
Repository federation training	In person	15 Staff + \$50K per year for workshops	7000 --> Removing cost per attendee and using standard average costs for each band as requested by costing consultants	Training for LMIC repository managers
Criteria refinement workshops	In person	10 staff + \$200k per year for workshops	7000 --> Removing cost per attendee and using standard average costs for each band as requested by costing consultants	Workshops to refine federation criteria

8. CITIZEN ENGAGEMENT: Is citizen engagement involved in delivering the solution?

Activity	Location	Estimate of total number of HOURS of citizen engagement during the length of the solution:	this figure should reflect the time that will be compensated – it will also be used to inform the level of costs of reimbursing expenses, including travel	OTHER REMARKS
Integrate repository with global platforms and policymaking tools		50 --> 360*5	Suggestion for increasing the number of hours dedicated to citizen engagement	10 hours per month per repo so 360 x 5 10hrs*12months*3repos* 5 years

9.TRAVEL: Is travel and accommodation required for the administration of this solution and/or its promotion at meetings, conferences etc

Expected # overseas trips during the length of the solution	Expected # in-country trips	Total days of accommodation	OTHER REMARKS
0	10	260	3 in-person meetings for 3
24	0	72	days each, in which the
			Secretariat will be

10.PUBLICATIONS AND PROMOTION EXPENSES: Are there any dissemination activities involved in delivering your solution?

Type	# of items	Approximate cost of each item if known (in US\$)	ANY OTHER REMARKS
Open access publications	3	\$2,000	
Presentation materials		\$1,000	
Posters			
Videos	1	\$2,000	
Fees for attending conferences	6	\$1,200	

11.OTHER: Is there any other resource use that is not covered above? If so, please provide a brief description and estimated cost.

Launch long-term funding and sustainability strategy
--

WG 2 Data Sharing and Reuse

2

Developing Standardized Record Structures and Interoperable Formats

PEOPLE: Who will be involved in delivering the solution: this might include a wide range of professional, technical and administrative staff, including people who might have roles as trainers						
Activity	Role	Type	Location	Expected Number of staff	For how many months?	% Approx FTE
Standards development (overall)	Research lead	Senior		1	60	100%
	Project manager	Senior		1	60	100%
Mapping existing standards & tools	Research Expert	Senior		1	60	100%
Data modelling (initial 6 months + 48 months of buildup)				3		
	Data architects	Senior			60	100%
	Advisors	Senior		6	60	10%
Formats development (initial 6 months + 48 months of buildup)				5		
	Research Fellows	Mid-tier			54	100%
	Advisors	Senior		3	60	
Piloting & refinement	Research assistants	Mid-tier		1	18	100%
Outreach						

Activity	Type of external service	Expected value of the contract (in US\$):

[illegible]

7. CAPACITY DEVELOPMENT: Are you proposing the solution includes funding to pay people (staff or others) to attend capacity building activities (e.g. short course/ training/CPD etc)?

CAPACITY DEVELOPMENT: Are you proposing to participate in any training activities (e.g. short courses, training, CPD, etc.)?		Location of capacity building activity		Over the length of the solution, expected # of attendees funded (staff or others as part of the solution) to participate in the capacity building events		Average cost per attendee in US\$ (including fees, accommodation, travel, etc.) if known	
Activity							
Pilot conversion workflows with selected institutions (e.g., LMIC partners, sectoral repositories)				175			
						At least 5 workshops with selected institutions in different countries. In each one of them, there will be 30 participants	
						training event for orgs & maintaining existing tools	

CITIZEN ENGAGEMENT: Is citizen engagement necessary to solving the solution?		Location	Estimate of total number of HOURS of citizen engagement during the length of the solution:	this figure should reflect the time that will be compensated – it will also be used to inform the level of costs of reimbursing expenses. including travel		
Activity						
Launch open consultation and finalize standards based on feedback			500		Citizens will be engaged during piloting workshops as well as during open consultation	Suggestion for increasing the number of hours dedicated to citizen engagement

Expected # overseas trips during the length of the solution	Expected # in-country trips	Total days of accommodation	
40	10	300	2 in-person meetings for 3 days each across the Secretariat and the advisory group.

Type	# of Items	Approximate cost of each item if known (in US\$)
OA papers covering core components	5	2000
Fees for attending conferences	6	\$800

11. OTHER: Is there any other resource use that is not covered above? If so, please provide a brief description and estimated cost.

1.Name of the Group: WG 2 Data Sharing and Reuse
2.Solution number in the order sent to the Costing team [use A, B etc for any 3
3.Title of the solution: Establishing Metadata standards to facilitate data identification and discoverability

4.PEOPLE: Who will be involved in delivering the solution: this might include a wide range of professional, technical and administrative staff, including people who might have roles as trainers

Activity	Role	Type	Location	Expected Number of staff	For how many months?	% Approx FTE
Standards development (overall)	Research lead	Senior		1	60	100%
	Project manager	Senior		1	60	100%
Mapping existing standards & tools	Research Expert	Senior		1	60	100%
Data modelling (initial 12 months + 48 months of build up)	Data architects	Senior		3	60	100%
	Advisors	Senior		6	60	10%
Formats development (initial 6 months + 48 months of build up)	Research Fellows	Mid-tier		5	54	100%
	Advisors	Senior		5	60	
Piloting & refinement	Research assistants	Mid-tier		1	18	100%
OUT reach + institutionalize metadata schema and QA practices across ESIC tools	Metadata Governance Lead	Mid-tier		1	60	100

5.Are there any third-party contracts (except, e.g. IT maintenance which is included in Q6) involved in delivering your solution, i.e. costs charged by an external supplier to deliver a piece of work? If so, please indicate:

Activity	Type of external service	Expected value of the contract (in US\$):
Software license		
Data management tools		
Data sharing tools		
AI / ML Subscriptions	Subscription	5000 usd / year
Multilingual Validation	Translation Services	40000

External translation for metadata fields

6.TOOLS: Would you need any tool to deliver this solution? If so indicate all that applies to you – this could be existing tools, or items required for developing a new tool.

Activity	Type	How many?	Approx cost of each in US\$ if known	For how many months where relevant?	OTHER REMARKS
Software license (
Data management tools					
Data sharing tools					Maybe can have opportunities for cost-sharing or collaboration with existing evidence-sharing platforms? NDC
Metadata schema	Ontology builders	1	2000/year	30	Tools for schema development
Tagging tools	Tagging engines	1	3000/year	24	AI-assisted tagging tools
	Hardware	2	Laptops for each staff member every 2 years (1 x person at employment and 1 x person upgrade during the 5 years, i.e. upgrade after 3-4 years)		

OTHER SYSTEM REQUIREMENTS

7.CAPACITY DEVELOPMENT: Are you proposing the solution includes funding to pay people (staff or others) to attend capacity building activities (e.g. short course/ training/CPO etc)?

Activity	Location of capacity building activity	Over the length of the solution, expected # of attendees funded (staff or others as part of the solution) to participate in the capacity building events	Average cost per attendee in US\$ (including fees, accommodation, travel, etc.) if known	OTHER REMARKS
Pilot tagging system across selected ESIC-aligned repositories		175	Using standard average costs for each band as requested by costing consultants	Several pilots that will entail at least 5 workshops with selected institutions in different countries. In each one of them, there will be 30 participants (+5 staff members). The workshop will have a practical session to test the typologies/taxonomies in practice
Train-the-trainer programs	Regional hubs			

8.CITIZEN ENGAGEMENT: Is citizen engagement involved in delivering the solution?

Activity	Location	Estimate of total number of HOURS of citizen engagement during the length of the solution:	This figure should reflect the time that will be compensated – it will also be used to inform the level of costs of reimbursing expenses, including travel	Citizens will be engaged during piloting workshops as well as during open consultation	Suggestion for increasing the number of hours dedicated to citizen engagement
Finalize standards through open consultation and feedback		500			

9.TRAVEL: Is travel and accommodation required for the administration of this solution and/or its promotion at meetings, conferences etc

Expected # overseas trips during the length of the solution	Expected # in-country trips	Total days of accommodation
30	10	240

2 in-person meetings for 3 days each

10.PUBLICATIONS AND PROMOTION EXPENSES: Are there any dissemination activities involved in delivering your solution?

Type	# of items	Approximate cost of each item if known (in US\$)

11.OTHER: Is there any other resource use that is not covered above? If so, please provide a brief description and estimated cost.

1.Name of the Group:

WG 2 Data Sharing and Reuse

2.Solution number in the order sent to the Costing team [use A, B etc for any

4

3.Title of the solution:

Setting Open Access Standards for Equitable Data Sharing and Reuse

4. PEOPLE: Who will be involved in delivering the solution: this might include a wide range of professional, technical and administrative staff, including people who might have roles as trainers

Activity	Role	Type	Location	Expected Number of staff	For how many months?	% Approx FTE	
Legal strategy	Legal counsel	Senior		1	60	100	Develops licensing standards
Partnerships leads	Advocacy team	Senior		2	60	100	
Engagement Leads	Advocacy team	Senior		2	60	100	
Policy Leads	Advocacy team	Senior		2	60	100	
Program management	Senior program manager	Senior		1	60	100	
Advisory group	Advisors	Senior		10	60	10	Support the process - Convene power
Comms management	Senior comms manager	Senior		1	60	100	
Comms management	Comms Officer	Mid Level		1	60	100	

5.Are there any third-party contracts (except, e.g. IT maintenance which is included in Q6) involved in delivering your solution, i.e. costs charged by an external supplier to deliver a piece of work? If so, please indicate:

Activity	Type of external service	Expected value of the contract (in US\$):	
Advertising open licensing / Communications	Advertisement	\$ 10,00,000	Comms budget and make it \$1m ie 200k per year
Developing a website for disseminating the solution	Website development	\$ 20,000	
Social Media channel channel to keep the community up to date with progress and new systems			
Public relations and advocacy	Consultancies in public relations	5000000	

6. TOOLS: Would you need any tool to deliver this solution? Pls indicate all that applies to you – this could be existing tools, or items required for developing a new tool.

Activity	Type	How many?	Approx cost of each in US\$ if known	For how many months where relevant?	OTHER REMARKS
	Hardware	2	Laptops for each staff member	60	

OTHER SYSTEM REQUIREMENTS

7. CAPACITY DEVELOPMENT: Are you proposing the solution includes funding to pay people (staff or others) to attend capacity building activities (e.g. short course/ training/CPD etc)?

Activity	Location of capacity building activity	Over the length of the solution, expected # of attendees funded (staff or others as part of the solution) to participate in the capacity building events	Average cost per attendee in US\$ (including fees, accommodation, travel, etc.) if known	OTHER REMARKS
Validation workshops for open-access concepts		175		At least 5 workshops with selected institutions in different countries. In each one of them, there will be 30 participants.

8. CITIZEN ENGAGEMENT: Is citizen engagement involved in delivering the solution?

Activity	Location	Estimate of total number of HOURS of citizen engagement during the length of the solution:	this figure should reflect the time that will be compensated – it will also be used to inform the level of costs of reimbursing expenses, including travel	OTHER REMARKS
Engaging citizens as part of the validation workshops		50 ->500		Citizens will be engaged during workshops Suggestion for increasing the number of hours dedicated to citizen engagement

9. TRAVEL: Is travel and accommodation required for the administration of this solution and/or its promotion at meetings, conferences etc

Expected # overseas trips during the length of the solution	Expected # in-country trips	Total days of accommodation	
30	10	240	2 in-person meetings for 3 days
	0	36	Additionally, participation in 6 conferences
12			

10. PUBLICATIONS AND PROMOTION EXPENSES: Are there any dissemination activities involved in delivering your solution?

Type	# of items	Approximate cost of each item if known (in US\$)	OTHER REMARKS
Open access publications	1	\$2, 000	
Fees for attending conferences	6	\$800	
Fees for promotional material at conferences: QR codes for reports printed, post printing fees, any other printed material (if needed)			
Support for webinar series - access to webinar software + communications team			
Connection to global partners to present on research			

11. OTHER: Is there any other resource use that is not covered above? If so, please provide a brief description and estimated cost.

Translation services during travel

1.Name of the Group: WG 2 Data Sharing and Reuse

2.Solution number in the order sent to the Costing team [use A, B etc for any 5

3.Title of the solution: Ensuring Quality and Monitoring of Data Sharing and Reuse Systems

4. PEOPLE: Who will be involved in delivering the solution: this might include a wide range of professional, technical and administrative staff, including people who might have roles as trainers							
Activity	Role	Type	Location	Expected Number of staff	For how many months?	% Approx FTE	OTHER REMARKS
Develop QA standards	QA system researcher	Senior		1	60	100%	
Pilot QA systems	QA Manager	Senior		6	60	100%	
Project management	Project Manager	Senior		1	60	100%	
Conducting QA systems	QA Assistants	Senior	50		60	10%	
Develop QA standards/tech component	Product manager	Senior	1		60	100	
Develop QA standards/tech component	Team lead	Senior	1		60	100	
Develop QA standards/tech component	Designer	Senior	1		60	100	
Develop QA standards/tech component	Engineers	Senior	3		60	100	
Develop QA standards/tech component	Data scientists	Senior	3		60	100	
Develop QA standards/tech component	Data engineers	Senior	2		60	100	
Develop QA standards/tech component	Infrastructure engineer	Senior	1		60	100	
Development M&E strategy	M&E specialist	Senior	1		60	100	

5.Are there any third-party contracts (except, e.g. IT maintenance which is included in Q6) involved in delivering your solution, i.e. costs charged by an external supplier to deliver a piece of work? If so, please indicate:

Activity	Type of external service	Expected value of the contract (in US\$):	ANY OTHER REMARKS
M&E framework development	Consulting services	1,00,000	
Conduct periodic system-wide quality and equity evaluations	Evaluation services	70000	

6.TOOLS: Would you need any tool to deliver this solution? Pls indicate all that applies to you – this could be existing tools, or items required for developing a new tool.

Activity	Type	How many?	Approx cost of each in US\$ if known	For how many months where relevant?	OTHER REMARKS
QA standards	Framework builders	1	2,000/year	36	
Monitoring dashboards	Dashboard tools	1	3,000/year	36	Do we distinguish between one-off development costs and ongoing maintenance costs? -NDC
	Hardware	2	Laptops for each staff member	60	

OTHER SYSTEM REQUIREMENTS

7.CAPACITY DEVELOPMENT: Are you proposing the solution includes funding to pay people (staff or others) to attend capacity building activities (e.g. short course/ training/CPD etc)?

Activity	Location of capacity building activity	Over the length of the solution, expected # of attendees funded (staff or others as part of the solution) to participate in the capacity building events	Average cost per attendee in US\$ (including fees, accommodation, travel, etc.) if known	OTHER REMARKS
Conduct participatory review sessions with LMIC and sectoral partners		30		One participatory review session in which different partners would participate to provide feedback about quality and M&E standards

8.CITIZEN ENGAGEMENT: Is citizen engagement involved in delivering the solution?

Activity	Location	Estimate of total number of HOURS of citizen engagement during the length of the solution:	this figure should reflect the time that will be compensated - It will also be used to inform the level of costs of reimbursing expenses, including travel	OTHER REMARKS
Conduct participatory review sessions with LMIC and sectoral partners		500		As part of the participatory review process, citizens will also be engaged Suggestion for increasing the number of hours dedicated to citizen engagement

9.TRAVEL: Is travel and accommodation required for the administration of this solution and/or its promotion at meetings, conferences etc

Expected # overseas trips during the length of the solution	Expected # in-country trips	Total days of accommodation	OTHER REMARKS
40	10	200	2 in-person meetings for 3 days each

10.PUBLICATIONS AND PROMOTION EXPENSES: Are there any dissemination activities involved in delivering your solution?

Type	# of items	Approximate cost of each item if known (in US\$)	OTHER REMARKS
Open access publications	3	\$2, 000	
Fees for attending conferences	10	\$800	

11.OTHER: Is there any other resource use that is not covered above? If so, please provide a brief description and estimated cost.